An Evaluation of Voice Stress Analysis Techniques In a Simulated AWACS Environment

William Archer Jones, Jr.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	August 1990	Thesis/~~Dissertation~~

4. TITLE AND SUBTITLE
An Evaluation of Voice Stress Analysis Techniques in a Simulated AWACS Environment

5. FUNDING NUMBERS

6. AUTHOR(S)
William Archer Jones, Jr.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
AFIT Student at: Texas A&M University

8. PERFORMING ORGANIZATION REPORT NUMBER
AFIT/CI/CIA - 90-115

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
AFIT/CI
Wright-Patterson AFB OH 45433

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for Public Release IAW AFR 190-1
Distribution Unlimited
ERNEST A. HAYGOOD, 1st Lt, USAF
Executive Officer, Civilian Institution Programs

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

15. NUMBER OF PAGES
59

16. PRICE CODE

20. LIMITATION OF ABSTRACT

UNCLASSIFIED

GENERAL INSTRUCTIONS FOR COMPLETING SF 298

The Report Documentation Page (RDP) is used in announcing and cataloging reports. It is important that this information be consistent with the rest of the report, particularly the cover and title page. Instructions for filling in each block of the form follow. It is important to ***stay within the lines*** to meet *optical scanning requirements*

Block 1. Agency Use Only *(Leave blank)*

Block 2. Report Date. Full publication date including day, month, and year, if available (e.g. 1 Jan 88). Must cite at least the year.

Block 3. Type of Report and Dates Covered. State whether report is interim, final, etc. If applicable, enter inclusive report dates (e.g. 10 Jun 87 - 30 Jun 88).

Block 4. Title and Subtitle. A title is taken from the part of the report that provides the most meaningful and complete information. When a report is prepared in more than one volume, repeat the primary title, add volume number, and include subtitle for the specific volume. On classified documents enter the title classification in parentheses.

Block 5. Funding Numbers. To include contract and grant numbers; may include program element number(s), project number(s), task number(s), and work unit number(s). Use the following labels:

C	- Contract	PR	- Project
G	- Grant	TA	- Task
PE	- Program Element	WU	- Work Unit Accession No.

Block 6. Author(s). Name(s) of person(s) responsible for writing the report, performing the research, or credited with the content of the report. If editor or compiler, this should follow the name(s).

Block 7. Performing Organization Name(s) and Address(es). [illegible]

Block 8. Performing Organization Report Number. [illegible]

Block 9. [illegible] Monitoring Agency Name(s) [illegible]

Block 10. [illegible] Monitoring Agency [illegible] (If known)

Block 11. [illegible]

Block 12a. Distribution/Availability Statement. Denotes public availability or limitations. Cite any availability to the public. Enter additional limitations or special markings in all capitals (e.g. NOFORN, REL, ITAR).

DOD	-	See DoDD 5230.24, "Distribution Statements on Technical Documents."
DOE	-	See authorities.
NASA	-	See Handbook NHB 2200.2.
NTIS	-	Leave blank.

Block 12b. Distribution Code.

DOD	-	Leave blank.
DOE	-	Enter DOE distribution categories from the Standard Distribution for Unclassified Scientific and Technical Reports.
NASA	-	Leave blank.
NTIS	-	Leave blank.

Block 13. Abstract. Include a brief *(Maximum 200 words)* factual summary of the most significant information contained in the report.

Block 14. Subject Terms. Keywords or phrases identifying major subjects in the report.

Block 15. Number of Pages. Enter the total number of pages.

Block 16. Price Code. Enter appropriate price code *(NTIS only)*.

Blocks 17. - 19. Security Classifications. Self-explanatory. Enter U.S. Security Classification in accordance with U.S. Security Regulations (i.e., UNCLASSIFIED). If form contains classified information, stamp classification on the top and bottom of the page.

Block 20. Limitation of Abstract. This block must be completed to assign a limitation to the abstract. Enter either UL (unlimited) or SAR (same as report). An entry in this block is necessary if the abstract is to be limited. If blank, the abstract is assumed to be unlimited.

AN EVALUATION OF
VOICE STRESS ANALYSIS TECHNIQUES
IN A SIMULATED AWACS ENVIRONMENT

A Thesis

by

WILLIAM ARCHER JONES, JR.

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

August 1990

Major Subject: Industrial Engineering

AN EVALUATION OF
VOICE STRESS ANALYSIS TECHNIQUES
IN A SIMULATED AWACS ENVIRONMENT

A Thesis

by

WILLIAM ARCHER JONES, JR.

Approved as to style and content by:

Jerome J. Congleton
(Chair of Committee)

R. Dale Huchingson
(Member)

Samuel G. Schiflett
(Member)

G.K.Bennet
(Head of Department)

August 1990

ABSTRACT

An Evaluation of Voice Stress Analysis Techniques
in a Simulated AWACS Environment. (August 1990)

William Archer Jones, Jr., B.S., U.S. Air Force Academy
Chair of Advisory Committee: Dr. Jerome J. Congleton

The purpose of this study was to determine if voice analysis algorithms are an effective measure of stress resulting from high workload. Fundamental frequency, frequency jitter, and amplitude shimmer algorithms were employed to measure the effects of stress in crewmember communications data in simulated AWACS mission scenarios.

Two independent workload measures were used to identify levels of stress: 1) a predictor model developed by the simulation author based upon scenario generated stimulus events, and 2) the duration of communication for each weapons director, representative of the individual's response to the induced stress. Between eight and eleven speech samples were analyzed for each of the sixteen Air Force officers who participated in the study.

Results identified fundamental frequency and frequency jitter as statistically significant vocal indicators of stress, while amplitude shimmer showed no signs of any significant relationship with workload or stress. Consistent with previous research, the frequency algorithm was identified as the most reliable measure. However, the results did not reveal a sensitive discrimination measure between levels of stress, but rather, did distinguish between the presence or absence of stress.

The results illustrate a significant relationship between fundamental frequency and the effects of stress and also a significant inverse relationship with jitter, though

less dramatic. Applied research in this area must investigate the predictive power, within subjects, of these measurement techniques. Thesis. [illegible]

ACKNOWLEDGEMENTS

I can't begin to thank all the people who invested their time and energy into me during my tenure here at A&M.

A hearty thanks to Dr. Jerome J. Congleton, committee chairman, for your vision in the summer of '88 and constant support since. Thank you Dr. R. Dale Huchingson, for the wealth of knowledge and experience you shared with me during my stay . A special thanks to Dr. Schiflett and my new family at the School of Aerospace Medicine, for taking care of me and for all the great memories. I want to mention you all by name but I don't have five extra pages. To Tom Doherty, thanks for the software and patience needed to accomplish the goal. A big thanks goes out to my Homey, Lt Slater. Tim, I couldn't have done it without you!

Eternal thanks to my family, who are, and always will be, pillars of strength and love. Thanks Pastor Jeff, Marianne, and the Church for the love and encouragement I've received. Between you all, I know there were prayers for me heard in Heaven that could move 1000 mountains. Terri, thank you, just for being you and giving me the TLC I needed to "keep the Faith".

Thank You, Jesus, for Your faithfulness and grace to guide and sustain me. Jesus said unto him, "Thou shalt love the Lord thy God with all thy heart, and with all thy soul, and with all thy mind. This is the first and great commandment. And the second is like unto it, Thou shalt love thy neighbour as thyself. On these two commandments hang all the law and the prophets". *St. Matthew 22: 37-40*

TABLE OF CONTENTS

LIST OF FIGURES

INTRODUCTION

Psychological stress has been defined by researchers as a psychological state of tension that is a response to a perceived threat, accompanied by a specific emotion such as fear, anxiety, or anger (Biers, 1984). Due to the debilitating effects of this phenomenon on human performance, a variety of studies have addressed the causes of psychological stress, its impact on performance tasks, and how to measure its effects. Discussion of this topic relies, in part, on having a logical connection between the concepts of psychological stress and workload.

Although there is no universal agreement among researchers as to what constitutes workload, three general components are widely accepted: 1) a "busyness" factor characterized by time stress, a high frequency of cognitive involvement which incorporates higher attention demand, 2) a complexity factor that involves more difficult tasks requiring greater mental load, and 3) a psychological or emotional factor directly related to an individual's ability to cope (Biers, 1984). Laboratory experimentation of high workload environments attempt to artificially induce these three characteristics through stimulus manipulation, but a quantifiable and consistent stressor has proven to be elusive.

Measures of operator workload have been under scrutiny for years without a specific technique recommended as the definitive measure of workload, due to the multi-dimensionality of stress inherent in an operational setting (Gartner and Murphy, 1976 and Roscoe, 1978). Ideally, a technique should be objective and quantifiable, simple to collect and analyze, reliable across individuals, and a non-obtrusive stress measure readily applicable to dynamic field environments (Schiflett and Loikith, 1980).

This thesis follows the form and style of *Human Factors*.

Studies involving subjective workload assessment (SWAT) are non-obtrusive; however, they are subjective in nature and time-intensive in evaluation (Shingledecker and Crabtree, 1982). Physiological measures of stress involving heart rate, blood pressure, electromyographic (EMG) techniques, and measurement of pupil dilation are all proven objective measurements of stress (Brenner and Shipp, 1982). The equipment necessary for metrics of this kind are obtrusive and interfere with operator performance. Another measure of stress, which is an objective component and is potentially non-obtrusive, involves verbal communication in conjunction with the completion of a task.

There is sufficient evidence that acoustic parameters of the human voice provide reliable indicators of psychological stress (Schafer and Rabiner, 1975; Branscomb, 1979; Cannings, 1979; Schiflett and Loikith, 1980; Brenner and Shipp, 1988; Ruiz, Legros, and Guell, 1990). The aerospace community has been active in research involving voice stress analysis in support of the pilot and other crewmembers due to the increasing demands placed upon them by the technological advances seen in modern cockpits. Man has literally become the limiting factor in aerospace weapon systems because of our inability to process large quantities of information in an efficient manner. The increased potential of voice stress analysis in comparison to other stress measures is threefold: 1) it provides a non-obtrusive measure of pilot workload, 2) voice indicators of stress may provide warning information of an individual's inability to perform a given task, thereby, aiding in aircraft accident investigations, and 3) to determine the effects of stress on speech signal characteristics in voice actuated control technologies (Biers, 1984).

Current research has focused on seven candidate voice measures which show the greatest promise of responding to psychological stress. The measures include: 1) fundamental frequency (pitch) -under stress, pitch of the voice increases, 2) frequency jitter - under stress, the variability of pitch in the voice (jitter) decreases, 3) amplitude (loudness) - under stress, amplitude increases, 4) amplitude shimmer - under stress, variability of amplitude in the voice (shimmer) decreases, 5) psychological stress evaluator (PSE) scores - under stress, computer-derived PSE scores increase, 6) speech rate - under stress, the rate of speech increases, and 7) energy distribution - under stress, the proportion of energy between 500-1000 Hz increases (Brenner and Shipp, 1988; Kuroda et al., 1976; Schiflett and Loikith, 1980; and Scherer, 1981). Amplitude has only limited application in environments with consistent control over both distance and angle of the recording device to the crewmember's mouth and thus, may prove impractical for operational settings. Equipment and software were not available to include the PSE, speech rate, and energy distribution techniques, therefore, the present study will incorporate fundamental frequency, frequency jitter, and amplitude shimmer only.

A literature review investigating the utility of these voice measures "in the field" reveals inconclusive evidence as to their validity (Biers, 1984). However, the nature of the evidence must be challenged. The vocal parameters themselves may be sound, but the adequacy of the stressors employed are in question. Previous research, conducted in clinical settings with voice and physiological parameters of stress, has shown positive correlation between these measures (Brenner and Branscomb, 1981; Brenner and Shipp, 1982; and Brenner et al, 1983).

This study employed a similar approach to evaluate voice stress measures involving non-obtrusive, independent criterion of psychological stress. Brenner and Shipp (1988) found several significant relationships between their chosen voice measures and artificially induced levels of stress. The present study utilized three of the same vocal measures but extended the research beyond simple synthetic laboratory tasks to a complex decision-making study, where situational stress was not artificially produced but the result of a realistic and dynamic military mission. Manipulation of this workload-intensive simulation and direct measurement of specified tasks within the simulation scenario provided the independent measures to effectively evaluate the chosen voice measures of stress.

The primary objective of this study was to evaluate and compare fundamental frequency, frequency jitter, and amplitude shimmer algorithms as voice stress measurement techniques in the analysis of crewmember communications data from a simulated Airborne Warning and Control System (AWACS) mission scenario. A regression and correlation analysis between the three dependent voice measurement techniques and two independent workload measures allowed for comparison with previous research results and provided insight as to potential applications in voice stress analysis technology.

METHODOLOGY

The methodology of this experiment was driven by the research protocol, "Comparative Effects of Antihistamines on Aircrew Under Sustained Operations", developed in accordance with USAF commitments to the Joint Working Group on Drug Dependent Degradation (JWGD3) of Military Performance (See protocol in Appendix A). This triservice project evaluated a nonsedative antihistamine medication terfenadine (Seldane) against a placebo condition and a positive control diphenhydramine (Benadryl), using complex Command, Control, and Communications (C^3) tasks and a variety of performance measures, including voice stress analysis. A second objective was to assess the performance degradation of individual and team performance produced by the C^3 scenarios of high and low workload (Eddy, 1989 and 1990).

Although the primary goals of the USAF project were drug related, the placebo and Seldane groups were analyzed in the present study without respect to different medication conditions. A repeated measures Analysis of Variance (ANOVA) was used to compare Medication X Time and no significant differences were found between the Seldane medication and placebo effects (C. Oakley, personal communication, May 1, 1990). Therefore, speech communications from both groups were used for the stress analysis data.

Subjects

Sixteen Air Force officers (thirteen male and three female with an average age of 26) from the 552d Air Wing, Tinker Air Force Base, Oklahoma, were assigned to Brooks Air Force Base, Texas, to spend their work week in support of the study. The subjects were all weapon directors (WDs) with between 265-2000 hours of flight and simulator time logged (See subject data in Appendix B). All participants had previously volunteered for the study. WDs are primarily responsible for C^3 missions in the airborne environment. The E-3A AWACS aircraft is the platform from which WDs accomplish their mission. All subjects signed the Advisory Committee on Human Experimentation (ACHE) approved consent form prior to any data collection and female subjects had a pregnancy test within the previous thirty days and signed a pregnancy disclaimer (Appendix A).

Apparatus

AESOP Facility. The Aircrew Evaluation Sustained Operations Performance (AESOP) facility provides a flexible and realistic operational environment that supports Department of Defense (DoD) research in a variety of aerospace and space applications. Integration of hardware and software resources produced seven empirically derived air defense AWACS scenarios. Laboratory configuration consisted of Command, Control, and Communications Generic Workstations (C^3GW) which provided realistic mission simulation, tasks, controls, and authentic graphic displays replicating the functions of the AWACS WD crewstation (Figure 1).

AESOP assets also consist of supporting VAX, MicroVAX III, and VAX Station computer systems, a 16-channel RACAL FM recorder, video cameras, and a 10-node audio communication network which provides audio communication during simulations. This research tool allows the military services to develop and transition performance enhancement techniques and methods of analysis from the laboratory to field test experiments, and eventually to actual military operations.

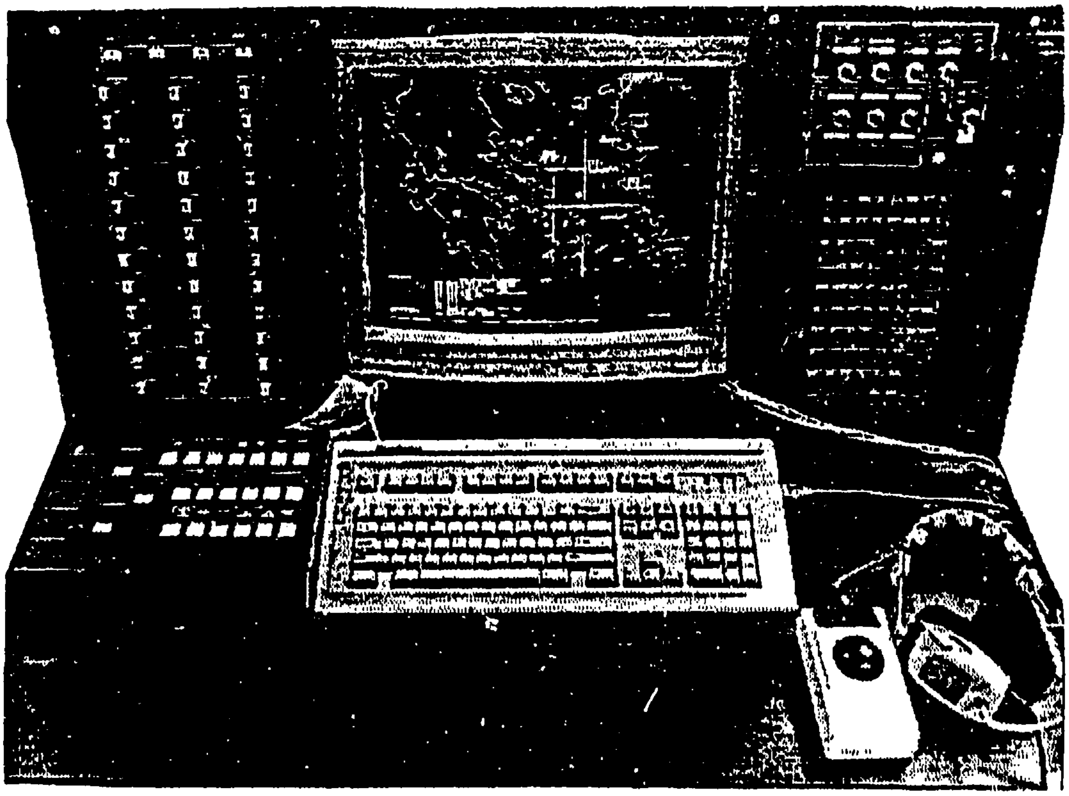

Figure 1. Laboratory configuration of the AWACS WD crewstation (C^3GW).

Simulation Scenarios. The simulation scenarios were designed to provide clearly challenging work environments in order to produce psychological and physiological differences within subjects tested. As the validity of voice stress effects had been a stumbling block in past research (Biers, 1984), the primary concern in development of these scenarios was to place realistic task demands on WDs and have an accurate system of measurement to capture performance data.

A Defensive Counter Air mission (DCA) was utilized because it encompassed the widest variety of WD performance tasks. The DCA mission consists of operations conducted to detect, identify, intercept, and destroy enemy aircraft attempting to penetrate friendly airspace. Each scenario was based on a standard enemy attack of four waves. Figure 2 is a model of the predicted workload throughout a three-hour scenario, measured in terms of WD activity units per five minute interval. Minutes 0-15 represent the baseline level area, while the four waves are depicted by the rise in activity units at minutes 30, 60, 90, and 120. Figure 3 is a graph of a typical individual WD's measured duration of communication per five minute interval.

Data Reduction Software. Analog-to-digital (A/D) and digital-to-analog (D/A) conversion software, as well as digital scoring algorithms of each stress metric were programs (written in TurboPascal 5.0) developed and provided by E. Thomas Doherty, Ph.D., of the Speech Research Laboratory, San Francisco, California. A Zenith 248 PC configured with a math coprocessor and a Data Translation A/D board # DT-2801A was used to run the analysis software. A Kron-hite filter created a bandpass of 20-3000 Hz through which all speech data traveled from the RACAL recorder to the computer A/D system. The filter eliminated frequency data in the signal that, in the present application, was considered "noise", to ensure clean and easily recognizable pitch periods for the analysis programs.

Four specific menu-driven programs were available for data reduction (Doherty, 1989). Targeted periods of communication were digitized from the analog tapes using the ADC software program. The parameters of 10,000 samples per second, for one channel, and a gain of four was used. ADC produced an output file read by the pitch extraction program, SWIFFT. This automated program inputs the digitized speech and displays it in an analog format on the screen. Still in the time domain, SWIFFT uses a

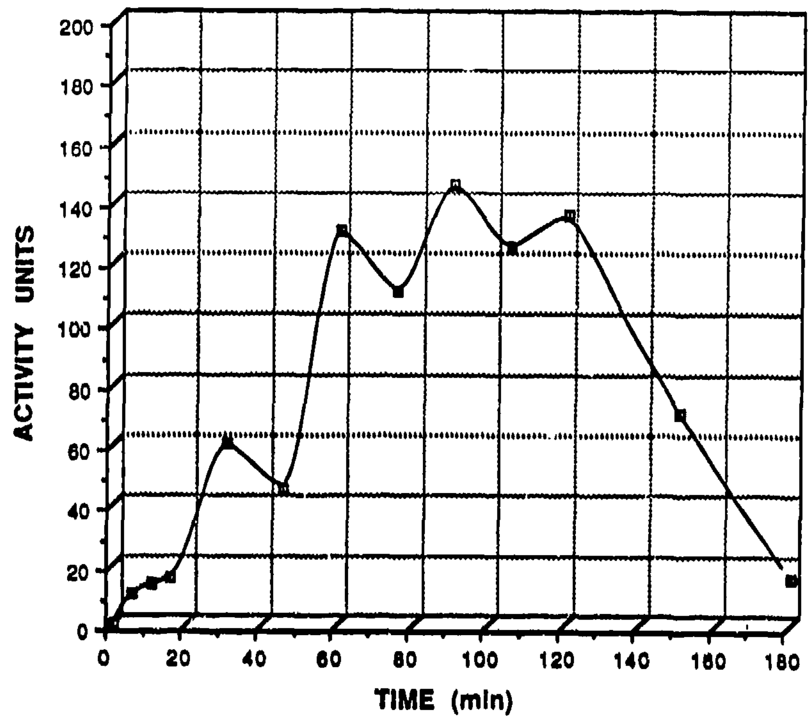

Figure 2. Graph of predictor model workload measurement.

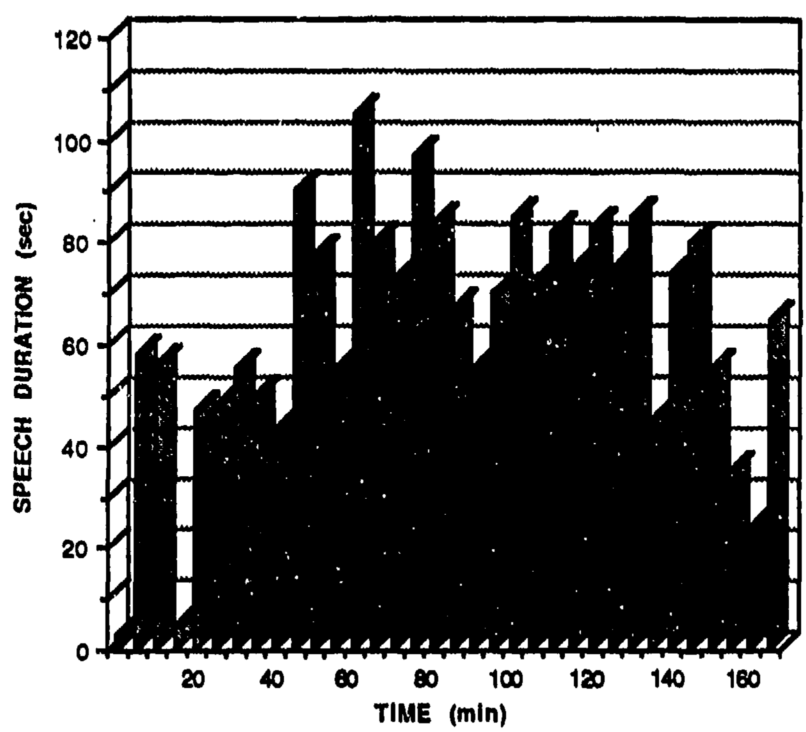

Figure 3. Graph of duration of communication workload measurement.

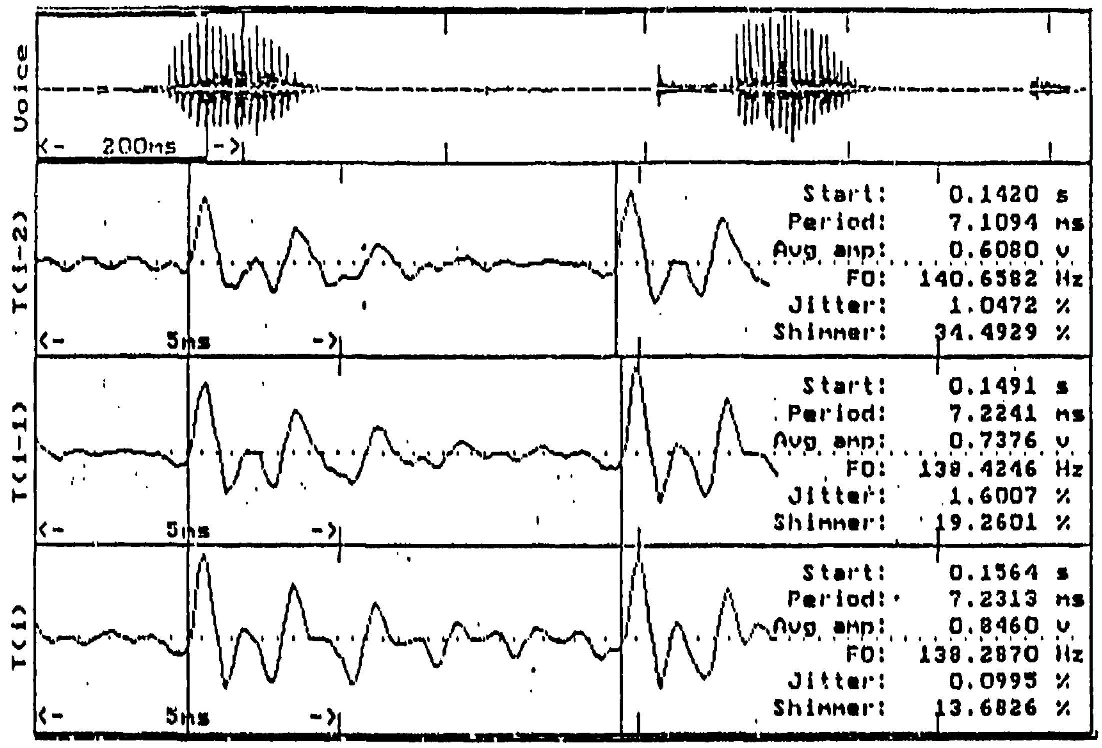

Figure 4. Optional graphic display during pitch extraction in SWIFFT. The wave form of each period may be shown to permit visual confirmation of correct markings of successive cycle boundaries.

peak-to-peak scoring algorithm to determine fundamental frequency, jitter, shimmer, and amplitude (Figure 4).

The F0CONV program reads the SWIFFT output and produces data files listing extraction values for each pitch period and a summary table of mean speech analysis values for that utterance (Summary data in Appendix C). TRUNCA2D is a program which allows the analyst to truncate just the relevant periods of speech so as to ease manipulation and storage of the digitized data.

Procedure

Subject WDs arrived at Brooks AFB on Sunday evening for a preliminary briefing. Training began on Monday, with a pre-brief beginning at 1130 hours and a three hour training scenario to familiarize them with the simulated AWACS crewstations and flight missions.

Teams were tested in two, three and a half hour scenarios per day on Tuesday, Wednesday, and Thursday. The more difficult simulation on Thursday, the Aaragon scenario, was chosen as the target simulation for this study because any confounding measures of stress due to the unfamiliar environment, new crewmembers, or learning curves were minimized by the third day of testing.

The 16-channel RACAL FM recorder taped all speech inputs through the Audio Distribution System (ADS) so the data could be collected for A/D conversion and analysis. After thorough review, sampled speech utterances were time-stamped and captured using the data reduction software. The human voice is an extremely dynamic environment to measure and the value of a vocal parameter can be content-dependent. Therefore, the selection of a consistent voiced utterance within subjects was imperative for accurate analysis. The voiced sequences of the words "Roger" and "Copy" were most often used, depending on the individual's most common vocabulary. Sampled utterances were most often targeted when attached to a string of words to negate any possible confounding affects from a single word.

Experimental Design

Individuals have their own unique voice print, much like fingerprints. The design of this experiment responded to the fact that the absolute values of the WD's voice parameters were of little value, when comparing between individuals. Data was analyzed in terms of relative changes in voice throughout the simulation profile, by utilizing within subjects analyses.

Two independent measures of stress were employed to identify times of high workload in the scenario. The predictor model provided simulation-dependent workload measures. The simulation author developed this model, based upon scenario manipulations such as number of hostile aircraft launched, using WD activity units per five minute interval as a measure of workload. Activity units are defined as the summation of 1) the number of footswitch microphone activations and 2) the number of tactical bearing and range, commit, and reinitiate button pushes. The duration metric provided individual WD-dependent data, based upon the crewmembers' reaction to scenario events. The WD's duration of communication within five minute intervals was considered an accurate measure since their primary duty is to communicate C^3 information between the AWACS and pilots under their control. The three voice measures of stress (fundamental frequency, jitter, and shimmer) were the dependent measures.

A two-way, repeated measures analysis of variance (ANOVA) model was used to evaluate the ability of the three voice algorithms to distinguish between five levels within the predictor model (baseline, low, medium, high-1, and high-2). Only the predictor model allowed for this analysis because the defined stress levels were consistent between subjects. Duncan's multiple range test was used for specific post-hoc comparisons between treatment means. A within subjects simple regression and correlation analysis was also employed to determine if a relationship exists (and to what extent) between the voice stress algorithms and each of the workload measures.

RESULTS

Predictor Model

The relationship between the five levels of workload (defined by the predictor model) as a function of frequency, jitter, and shimmer means is depicted in Figure 5.

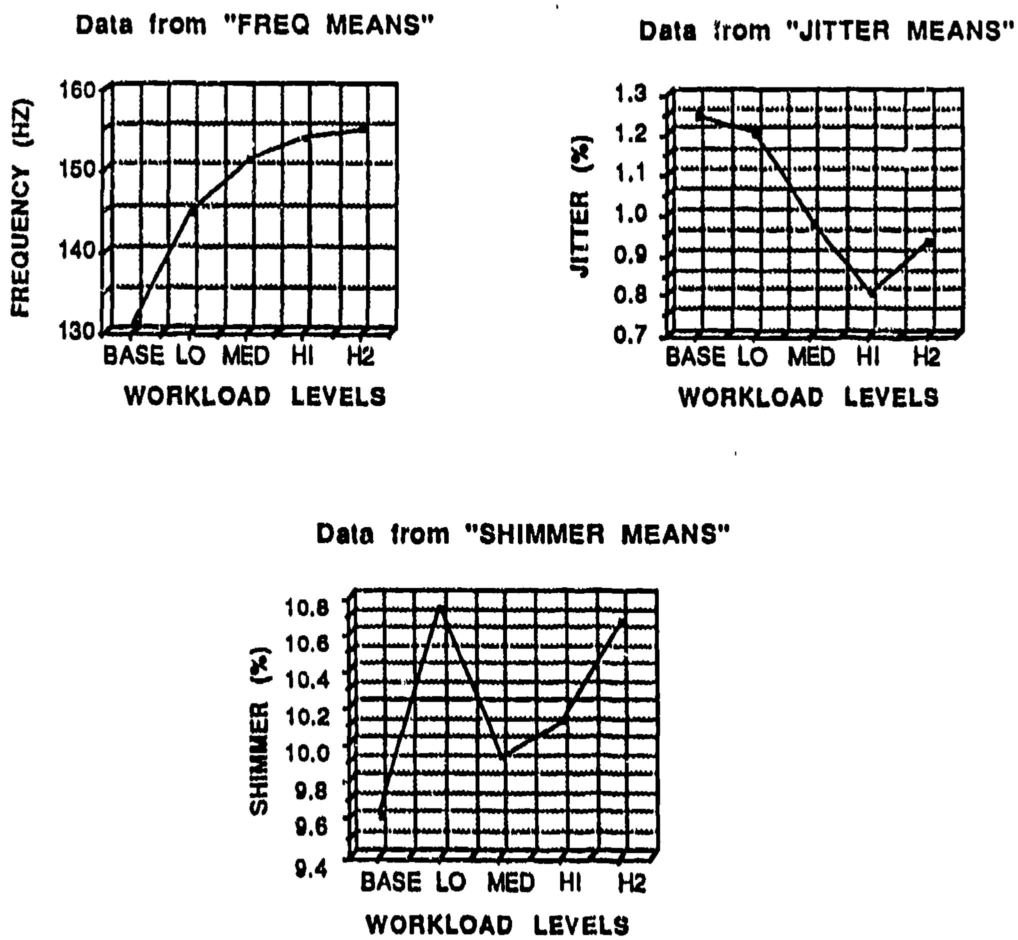

Figure 5. Predictor model analysis means as a function of frequency (upper left), jitter (upper right), and shimmer (bottom) at five levels of workload.

An ANOVA revealed that significant effects were obtained with the fundamental frequency measure, $F(4,15) = 7.63, p<.0001$. Follow-up Duncan's multiple range tests of the data show the baseline grouping to be significantly different from the four other levels of stress at ($p<.01$) and different from the med, h1, h2 levels at ($p<.001$). No other statistically significant differences exist between the workload levels. Both frequency jitter and amplitude shimmer parameters of voice were found to have no significant differences (See ANOVA tables in Appendix D).

Simple linear regression analysis was run on each WD (within subjects) to study the relationship between the three dependent variables and the predictor model (Appendix D includes WD regression and correlation values). The results showed the frequency measurement to be positively correlated with levels of stress in 15 of the 16 subjects (Figure 6). The pooled correlation was significant at the ($p<.0001$) level (See Table 1).

Table 1.

Pooled regression and correlation data.

Measure	N	Slope	P-Value	R-Value
Model Frequency	80	.163390	.0001	.616
Model Jitter	80	-.002909	.0245	-.323
Model Shimmer	80	.002996	.6406	.103
Duration Frequency	153	.271899	.0001	.591
Duration Jitter	153	-.006140	.0171	-.214
Duration Shimmer	153	-.016453	.3424	-.124

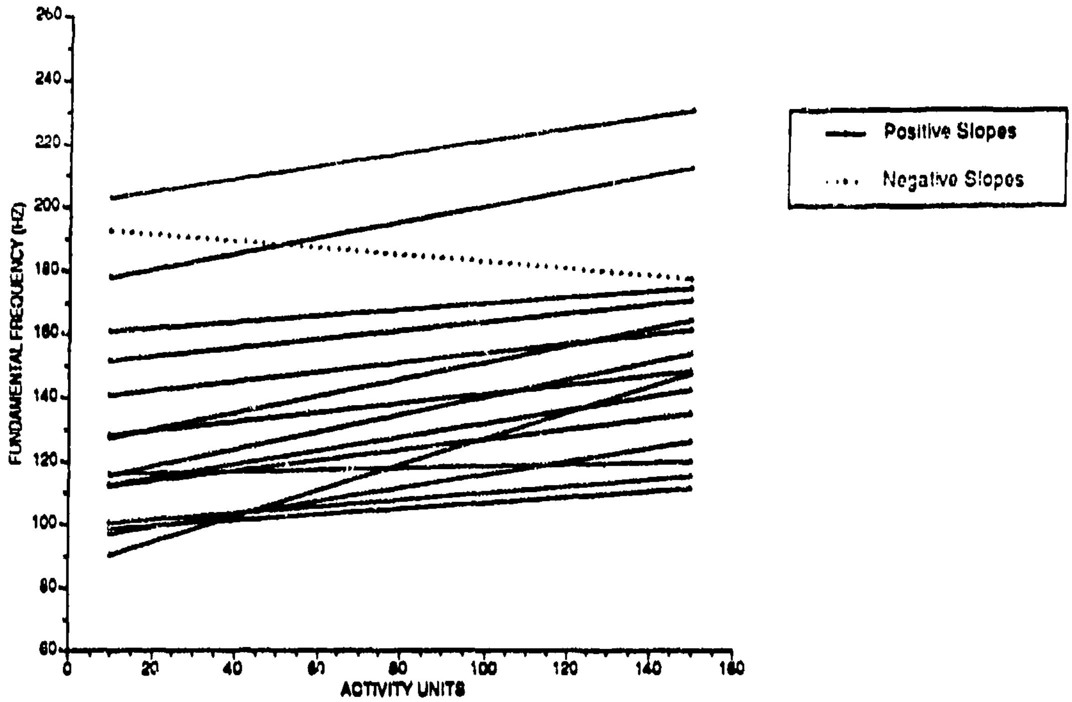

Figure 6. Predictor model regression slopes of activity vs frequency.

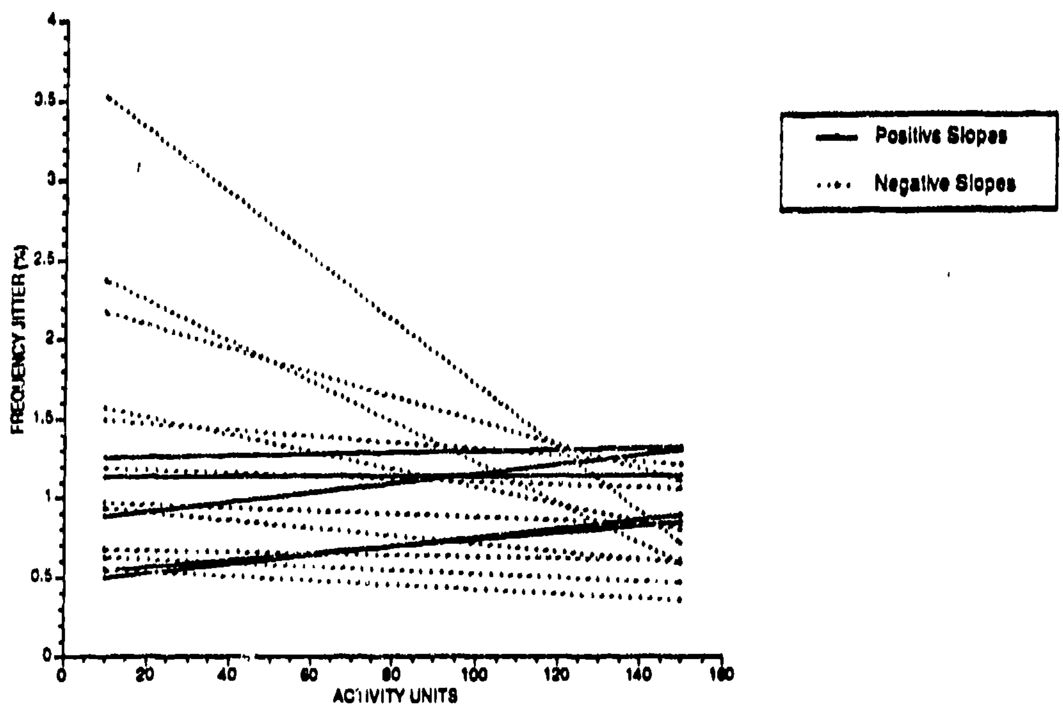

Figure 7. Predictor model regression slopes of activity vs jitter.

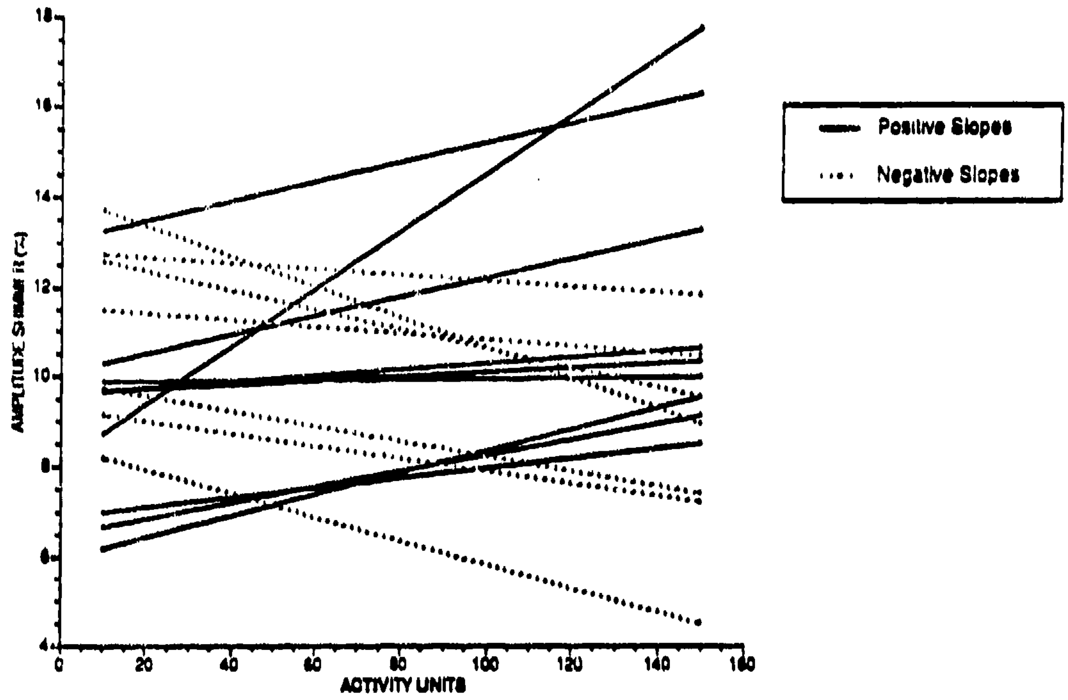

Figure 8. Predictor model regression slopes of activity vs shimmer.

A similar, though not as obvious, relationship was found to exist in the analysis of jitter, with 11 of the 16 negative slopes (Figure 7). Pooled data showed a significantly negative relationship with levels of stress ($p<.05$). Shimmer data revealed no significant relationship and approximately half of the slopes were found to be in a positive direction (Figure 8).

Duration of Communication

Total time involved in subject communication varied from 4-147 seconds within each five minute interval. The results of the duration of communication data analysis was similar to the model data results, but with a higher degree of confidence. The number of speech samples were increased per WD and the fundamental frequency measure showed a significantly ($p<.0001$) positive relationship, with all slopes following an upward trend (Figure 9). Jitter values were also shown to be significantly correlated ($p<.05$), but in the negative direction, as shown in Figure 10. Once again, the shimmer measure has shown no significant negative relationship with the duration of communication workload metric (Figure 11).

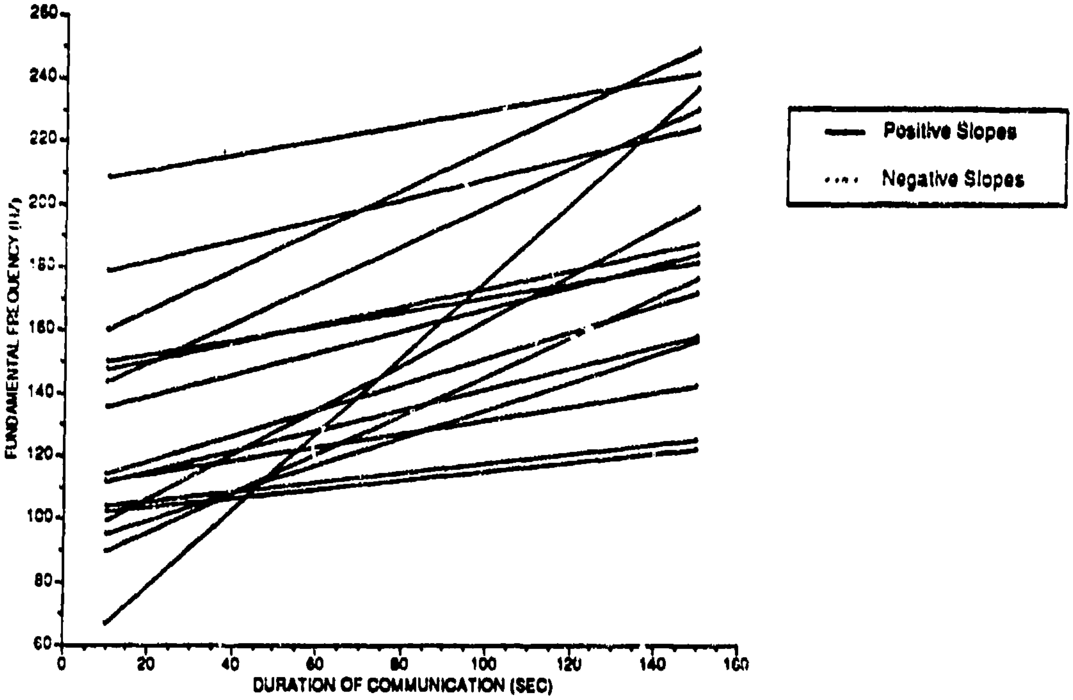

Figure 9. Duration of communication regression slopes of seconds vs frequency .

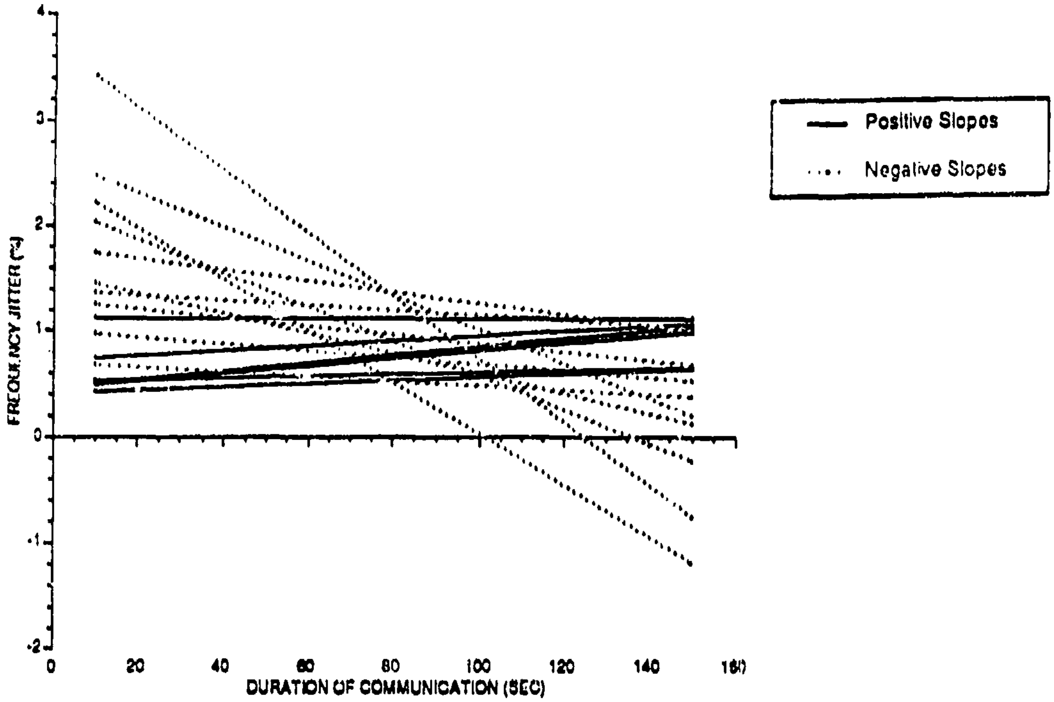

Figure 10. Duration of communication regression slopes of seconds vs jitter.

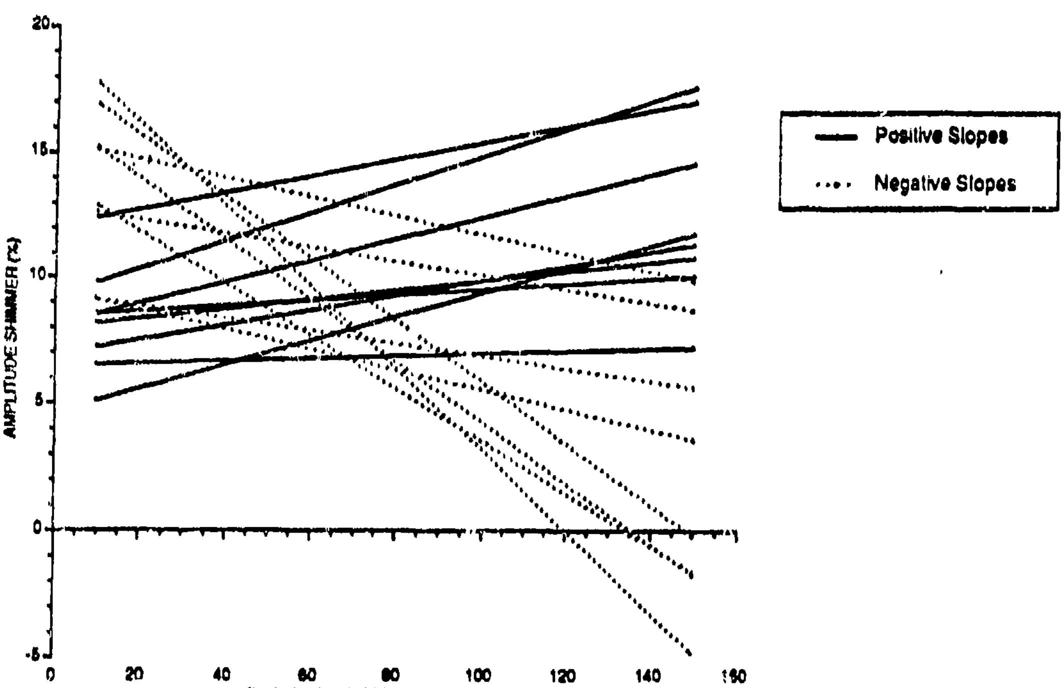

Figure 11. Duration of communication regression slopes of seconds vs shimmer.

DISCUSSION

The voice stress measurement results showed a consistent pattern with both the simulation-dependent predictor model of workload and the individual WD-dependent metric. The findings of this study are similar to previous results found by Brenner and Shipp (1988). For example, fundamental frequency proves to be the most salient vocal indicator of psychological stress. The frequency jitter measure responded in the predicted fashion, but with marginal statistical significance and less confidence. Amplitude shimmer produced insignificant results and showed no promising trends of a potential relationship with the effects of stress.

In the ANOVA, the variabilities in pitch and amplitude were found to have no significance, but highly significant fundamental frequency results indicate the measure's ability to recognize a difference in levels of stress, even though its sensitivity to discriminate between the five levels was found to be unsatisfactory. In operational terms, this study reveals that an individual's vocal pitch can distinguish between whether the WD is pre-flighting his system on the ground vs flying, but not whether he is in combat or controlling for a refueling mission. However, this does not mean that this voice measure has no potential operational application. Perhaps these findings are a result of the model representations of medium, h1, and h2 being too close together. The disparity between the medium level of activity and the h2 level is only 15 units. The model's peaks were selected for analysis to coincide with the scenario's high workload stimuli based on engagement of enemy aircraft waves.

The predictor model represented high task-loading events from the scenario but not every WD responds at the same time or in the same way to a given stimulus in the simulation. The crewmembers work in teams of three and an individual WD's reaction to the model's representation of high workload is a direct result of team interaction.

The team may be running behind on a mission and the actual stress effect may not be confronted until five or ten minutes after the scenario stimulus was engaged; and that loading may last throughout the initiation of the next stimulus. The fact that high task-loading occurred in the scenario is not the question; it is whether the predictor model identified the correct sequence of time related to subject speech. It may have occurred earlier or later or may not have been manifested in the words "Roger" or "Copy".

Regression analysis presented a positive relationship between the predictor model and the frequency measure and a negative slope relationship with the jitter technique. As displayed in Figures 6 and 7, the trend is extremely strong with frequency but jitter shows a much less consistent trend. Shimmer reflects no consistent trend in either direction as is evident in Figure 8, when a negative slope was the expected result.

The increased number of speech samples per WD and subject specific nature of the duration workload measure, resulted in greater statistical evidence of a positively correlated relationship with frequency and a negatively correlated one with jitter. Figure 9 exhibits a consistent trend with every regression line having a positive slope. Jitter has some regression lines that are not in the expected direction (Figure 10), so confidence of this result is not as strong. The shimmer measure shows no degree of sensitivity to the chosen voice measures. Figure 11 shows the random spread of slopes in both directions.

The duration of communication workload measure was more representative of the individual effects of stress that WDs experienced during the simulation, since it was empirically derived. Task-loading variance between subjects was taken into account in this measure, but the adopted definition of workload could be questioned. The WD's primary duty is vocal communication, however, during a five minute interval of high communication, the person could conceivably become stressed and unstressed several times. Decreasing the intervals of time would identify more distinct high workload areas but finding similar utterances throughout the scenario within such small periods would prove impossible in the present application.

CONCLUSIONS/RECOMMENDATIONS

This study utilized fundamental frequency, frequency jitter, and amplitude shimmer algorithms as voice stress measurement techniques in the analysis of crewmember communications data from a simulated AWACS mission scenario. The results of the study suggest that a consistent relationship exists between the frequency measure and levels of stress, while a less consistent inverse relationship exists between the jitter measurement and stress. Shimmer measurements reveal no consistent trends with high workload or stress.

The predictive correlational value of the results is not strong, but that was not the intent. Voice characteristics vary greatly among individuals. One can not expect an average positive slope correlated with workload, regardless of how many subjects run, to hold accurate predictive information as to an individual's voice response. What this data does show, is a confident relationship between fundamental frequency and high workload. This relationship substantiates the need for individual voice mapping to identify one person's characteristics to provide a predictive model of stress from the frequency measure.

The results of this study seem to follow past research and has identified fundamental frequency as the best vocal indicator of stress. Many are skeptical as to the potential of vocal indicators of stress and their application to operational environments. These parameters of speech may be undetectable in human conversation, but there is a relationship that seems to exist with possible application to the space and aerospace environment. Evaluating the different methods of measuring voice stress allows a more accurate analysis of workload-induced stress and greater flexibility in the application of stress analysis. These results not only provide data essential to the improvement of operational efficiency, but more importantly, the opportunity to preserve human lives through the notification of potentially hazardous conditions.

Future studies need to recognize the necessity for definitive independent stress measures to allow more conclusive evidence involving research of vocal indicators of stress. Objective physiological data is one accurate mode of measurement but due to the required fidelity of this AWACS simulation, all obtrusive measuring devices were not included. Also, this study was limited to only the raw individual performance measures collected from the AWACS data. Processed individual performance data, that will be provided by researchers at Brooks AFB in the near future, should increase the accuracy and precision of identifying specific workload occurrences. The need for accurate assessment techniques of specific times that individuals are stressed, can not be emphasized enough.

Continued research in this area is required to provide reliable measures of workload and stress. Detailed performance data from the Brooks AFB AWACS study should be used as independent workload measures to evaluate these and other methods of voice stress analysis. Also physiological measurement techniques should be employed whenever possible to allow accurate evidence of subject stress. The predictive power of these voice indicators, within subjects, is worthy of closer investigation. Greater samples per subject, to establish a regression line may prove to be a predictive tool as to that subject's workload and stress level.

This research confirms previous claims as to the potential for speech measurement as a non-obtrusive assessment of pilot workload and stress. Its applications extend from aiding in the investigation of aircraft accidents and mishaps to forecasting pilot disfunction caused by overload to providing filter algorithms to modify crewmember speech input in sensitive speaker-dependent voice recognition systems.

Voice stress analysis is maturing as a research area and promises a bright future in manned aerospace and space technologies.

REFERENCES

Biers, D. W. (1984). Voice stress as a workload measure: Review of the literature and recommendations for further study. Final report, Contract NOO421-83-D-0027, Naval Air Test Center, Patuxent River, Maryland.

Branscomb, H. H. (1979). The development of a measure for task-induced stress in speech. Unpublished master's thesis, Massachusetts Institute of Technology.

Brenner, M. and Shipp, T. (1982). Acoustics and physiological manifestations of psychological stress. Paper presented at the meeting of the Acoustical Society of America, Chicago.

Brenner, M., Shipp, T., Doherty, T., and Morrissey, P. (1983). Vocal measures of psychological stress: Laboratory and field data. Paper presented at the Conference on Physiology and Biophysics of Voice, University of Iowa. (Private collection, W. Jones).

Brenner, M. and Shipp, T. (1988). Voice stress analysis. Paper presented at 1987 Mental-State Estimation Workshop, Williamsburg, Virginia. (Private collection, W. Jones).

Cannings, R. (1979). Speech patterns and aircraft workload. In: survey of methods to assess workload. NATO AGARDOGRAPH No. 246, 155-129.

Doherty, E. T. (1989, October). Computers in voice clinics and voice laboratories. Paper presented at the Pacific Voice Conference, California. (Private collection, W. Jones).

Eddy, D. R. (1989). Selected team performance measures in the C^3 environment. USAFSAM-TR-87-25.

Eddy, D. R. (1990). Performance based on measures of individual and team complex decision-making. Paper presented at the 7th Command and Control Decision Aiding Conference, Dayton, Ohio. (Private collection, W. Jones).

Gartner, W. B. and Murphy, M. R. (1976). Pilot workload and fatigue: A critical survey of concepts and assessment techniques. NASA/Ames, NASA TN D-8365.

Kuroda, I., Fujiwara, O., Okamura, N., and Utsuki, N. (1976). Method for determining pilot stress through analysis of voice communication. Aviation, space, and environmental medicine, 47, 528-533.

Ruiz, R., Legros, C., and Guell, A. (1990). Voice analysis to predict the psychological or physical state of a speaker. Aviation, space, and environmental medicine, 266-271.

Schafer, R. W. and Rabiner, L. R. (1975). Parametric representation of speech. In D. Reddy (Ed.), Speech recognition. New York: Academic Press. 99-150.

Scherer, K. R. (1981). Speech and emotion states. In J. K. Darby (Ed.), Speech evaluation in psychiatry. New York: Grune and Stratton. 189-220.

Scherer, K. R. (1982). Methods of research in vocal communication: Paradigms and parameters. In K. R. Scherer and P. Ekman (Eds.), Handbook of methods in nonverbal behavioral research. Cambridge: Cambridge University Press. 136-198.

Schiflett, S. B. and Loikith, G. J. (1980). Voice stress analysis as a measure of operator workload. Patuxent River, Maryland: Naval Air Test Center, TM-79-334.

Shingledecker, C. A. and Crabtree, M. S. (1982). Subsidiary radio communication tasks for workload assessment in R&D simulations: II. Task sensitivity evaluation. Wright-Patterson AFB, Ohio: Air Force Aerospace Medical Research Laboratory, AFAMRL-TR-82-57.

APPENDIX A

BROOKS AFB RESEARCH PROTOCOL

RESEARCH PROTOCOL

1. Title: Comparative effects of antihistamines on aircrew under sustained operations.

2. Project/Task/Work Unit: 272911A8.

3. Principal Investigator/Organization/Telephone: Samuel G. Schiflett, Ph.D., USAFSAM/VNB/3464; Douglas R. Eddy, Ph.D., NTI, Incorporated/532-5723.

4. Associate Investigators/Organization: Jonathan French, Ph.D., USAFSAM/VNB/3464; Clark Shingledecker, Ph.D, NTI, Incorporated/(513) 254-3171.

5. Medical Consultant: Chief of Flight Medicine, USAF Clinic, Brooks AFB.

6. Statistician: Joseph R. Fischer, Jr., USAFSAM/VNS/3811.

7. Objective: The objective of this investigation is to evaluate the sensitivity of selected C^3 and synthetic performance measures to the effects of two antihistamine medications, Benadryl and Seldane. The measures include mission, system, individual, and team performance; embedded task workload; situation awareness; and synthetic cognitive and psychomotor predictors. A second objective is to assess the magnitude of individual and team performance impairment produced by the antihistamines during high- and low-workload C^3 scenarios.

8. Background: The triservice Joint Working Group of Drug-Dependent Degradation (JWGD3) on military performance is determining the impact of certain classes of drugs on the performance of aircrews solving a range of mission-related tasks in stressful environments. One area of interest involves the effects of antihistamines on complex C^3 decision-making performance by teams during sustained operations. Because of the drowsiness side effects, USAF aircrew personnel are grounded while taking centrally acting antihistamines for seasonal allergies or non-allergic rhinitis symptoms. This results in frequent interruption of flying schedules, loss of training, and disruption of crew rest schedules for non-symptomatic crew members, especially during sustained operations. Recently, however, new antihistamines purporting to have no drowsiness side effects have become available to AF flight surgeons.

Seldane (Terfenadine) is a noncentrally-acting, H-1 type antihistamine with non-sedating properties (Boggs, 1987). Benadryl (Diphenhydramine) is also an H-1 type antihistamine, but often produces a sedative effect due to direct central nervous system (CNS) activation (Spector, 1987). Seldane has shown little or no performance impairment when compared to the significant performance impairments shown with centrally-active antihistamines such as Benadryl (Fink & Irwin, 1979; Clarke & Nicholson, 1978; Nicholson, Smith, & Spencer, 1982; Nicholson & Stone, 1982; Kulshrestha, Gupta, Turner, & Wadsworth, 1978; Betts, Markman, Debenham, Mortiboy & KcKevitt, 1984).

All of the studies cited above used simple performance tasks. The impact of the newer terfenadine medication on complex tasks is unknown. Demonstration of an absence of adverse effects on complex tasks under the terfenadine condition could potentially reduce grounding time by supporting a medical flying waiver. Complex performance tasks, such as the Complex Cognitive Assessment Battery, are beginning to appear (Samet, Marshall-Mies & Albarian, 1987). They have not, however, been normed or validated against complex, real-world work environments. At the present time, assessing the performance effects of these medications can best be accomplished in a simulation of real-world complex tasks under sustained operations.

Another issue of importance to this study is the assessment of drug effects on teams solving complex problems. Although there are several models for evaluating teams, most require inputs from trained observers making subjective ratings. Reliable detection of subtle effects of medications and fatigue require objective, repeatable measures. Eddy (1989) and Dyer (1986) both concluded that no one has systematically developed and empirically tested a comprehensive theory of team performance. As a result, Eddy and Shingledecker, under contract to the Air Force, have developed a hierarchical performance assessment system. This system includes four levels of metrics that can be interrelated. The levels include measures of mission effectiveness, system performance, human performance, and performance capability. The upper levels were developed in conjunction with subject matter experts (SMEs) in AWACS C^3 tasks. The system and human performance levels include several types of objective team measures such as situational awareness, cooperation, cohesiveness, adaptation, and distribution of work.

The proposed study will use six empirically derived, unclassified, air defense AWACS scenarios to evaluate two antihistamine medications against a placebo using a wide variety of performance measures. Three of the scenarios are high workload and three are low, as verified by subject matter experts and pilot test teams.

9. Experimental Method:

a. Study Design: The proposed study will use a double-blind design with a different drug administered to each of three groups. The three drugs include Seldane, Benadryl, and placebo control. Teams of three subjects will be tested together under placebo and one drug (see Figure 1) in both high- and low-workload conditions over three days (see Figure 2). Each team will receive a different order of workload so as to balance the order of these treatments during the morning and early evening sessions (see Figure 2). Teams will be randomly assigned without replacement to one of these orders. Figure 3 shows the daily schedule of testing activities. It was necessary to include a placebo condition during the first testing day to ensure the performance equivalence of the three groups. Accordingly, this testing day will be single-blind in that the experimenters will be aware of the drug condition on this day only. The subjects will remain blind to the drug condition beginning the evening of the training day and continuing throughout the study.

Both simple cognitive tests (Perez, Masline, et al., 1987) and complex laboratory tasks of the CCAB will be administered each day between the C^3 scenarios. Correlations of these test scores with those of the tasks embedded into the scenarios will provide data to assess the feasibility of predicting complex "real-world" performance from laboratory tasks under the same medications.

We anticipate that the performance, reserve capacity, and subjective measures will show degradation with the Benadryl antihistamine, with fatigue (order effect), and with high workload, when compared with placebo. We do not anticipate any degradation of performance with the Seldane antihistamine.

b. Assignment: The 552d Air Wing will assign twelve teams of three weapons directors (subjects may be male or female), who have previously volunteered, to Brooks AFB to spend their work week in support of this study. Teams will be randomly assigned to a drug treatment condition and will receive one of the workload orders shown in Figure 2. Female subjects must have a negative pregnancy test within the previous 30 days and sign a pregnancy disclaimer.

c. Interventions: Subjects will arrive at Brooks AFB on Saturday or Sunday evening for a preliminary briefing. Female subjects without proof of a negative pregnancy test will be tested. Training will take place on Monday for approximately eight hours. Teams will receive training on six simple computerized tests and two complex tests over approximately four hours. They will also run a three-hour C^3 training scenario to familiarize them with the simulated AWACS crewstations and scenarios; no drugs will be administered. Subjects will ingest one Benadryl and one Seldane placebo at 2230 or prior to going to sleep.

Teams will be tested in two 3 hour scenarios each day for three days starting on Tuesday. Figure 3 shows an event time-line of the dose administration and experimental event schedule for each 24-hr session. Each group will ingest placebos only during the testing schedule for Tuesday. A randomly assigned team will ingest the recommended therapeutic dose of either Benadryl plus lactose placebo, Seldane plus lactose placebo, or both lactose placebo preparations starting on Tuesday evening. Total antihistamine/ placebo ingestion for each group will consist of either eight 50mg Benadryl and ten placebo preparations; four 60mg Seldane and fourteen placebo preparations; or eighteen placebo preparations.

In order to keep the experiment doubly-blinded, dosing regimens for all groups will follow that for both Benadryl and Seldane. Benedryl and Seldane will have different appearances, hence the concurrent schedules under all test conditions. Each medication and its placebo will look identical to prevent the identification of the drug by appearance. Therefore, each subject regardless of group will consume 18 capsules.

Before retiring on Monday at 2230, and again on Tuesday at 0600 after a normal breakfast meal, each team member will ingest two lactose placebos (see Figure 1). Before being driven to the Aircrew Evaluation Sustained Operations Performance (AESOP) facility, each team member will complete a sleep survey, a profile of moods survey (POMS), a subjective fatigue rating scale, and a symptom (side effects) questionnaire. At 0630 the senior director will present a normal AWACS briefing of the upcoming scenario. Next, the team will begin the first scenario for the day, which may be either high- or low-workload, depending on the order of their assignment (see Figure 2). Dosing with placebo will continue QID throughout the first day of testing, refer to Figure 3, until the fourth or evening dose, which will be either drug or placebo. Depending on their group assignment, subjects will ingest either Seldane (60mg) plus Benadryl placebo, Benadryl (50mg) plus Seldane placebo, or both lactose placebos before going to sleep at 2230 Tuesday evening.

At 0600 Wednesday, depending on their group assignment, each team member will ingest the second dose of their drug treatment with the other placebo (or two placebos) after their normal breakfast meal. All events of Tuesday will be repeated on Wednesday and Thursday. They will not take a drug or placebo Thursday at 2230.

Breakfast, lunch, and a snack will be provided as scheduled during the 24 hr session. A fatigue/mood/symptom questionnaire will be completed on each of those occasions. The cognitive performance task batteries will be presented once each day between the two scenarios (see Figure 3). The only free time will be in the evenings and subjects will be expected to eat light, keep their blood alcohol levels below (0.1%), and retire by 2230.

Caffeine intake will be restricted throughout the testing session. Decaffeinated sodas, herbal tea, and water will be available periodically during the off-task time. Smoking will be allowed in designated, outside areas, during off-task periods only. Meals will be as low fat and low protein as possible to prevent the slower absorption of drug into tissue due to plasma protein binding.

d. Statistics: Data collected will be evaluated using an ANOVA with two repeated measures (workload and day) on two grouping factors (drug groups and order of workload). Measures of team performance and team processes will be evaluated using a similiar statistical model, with the understanding that there will be fewer degrees of freedom. Nonparametric tests will be conducted with the mood/symptom questionnaire data. Multiple linear regression techniques will provide results concerning the predictive power of simple task measures to complex real-world measures under the various drug conditions.

10. Requirements for Human Subjects:

a. Hazards: The potential discomforts and risks of the experiment are listed below:

1. Terfenadine (Seldane): A selective antihistamine taken to alleviate the symptoms of seasonal allergies. The most frequently reported adverse side effect has been drowsiness, with other less frequent reports of headache, fatigue, dizziness, nausea, and dryness of the mouth/nose/throat. All side effects listed were, however, no different in their frequency of report from the placebo group tested (Sorkin and Heel, 1985).

2. Diphenhydramine (Benadryl): Similar side effects have been reported in the Physicians Desk Reference (PDR) for this antihistamine. However, one clinical study reviewed, found reports of weakness, difficulty focusing vision, and dry taste in mouth as being significantly higher in frequency than the placebo group (Fink and Irwin, 1979).

3. Fatigue from a working three long days may manifest itself as decreased attention to details, a lack of interest or motivation, an increase in minor physical

complaints, substandard performance, or irritability (USAF Physiological Training Manual, AFP-160-5, 11-7, 1976).

The potential for experiencing all these side effects is not likely. With the symptoms questionnaire completed during the three meals each day, the investigators will have information concerning the incidence of such side effects for medical consultation, if necessary.

b. Benefits: These subjects will be on normal active duty and will receive no additional monetary benefits.

c. Source: TAC headquarters will assign all subjects to this study from a group who have previously volunteered. Subjects will be screened for prior prescription medication, including antihistamine use and reactions to lactose. Subjects currently on medication, including antihistamines will be excluded from this study.

d. Privacy: All provisions of the Federal Privacy Act, 5 USC 552a, will be complied with.

e. Stopping Rule: The subject may at any time elect to discontinue participation. The Medical Monitor may also stop a subject from participating or break the drug treatment code. The Command, Control Center will have a copy of the code that can be broken if a problem arises after normal clinic duty hours.

11. Support personnel/Instrumentation/Equipment: Utilization of the AESOP facility and all support staff is a requirement of this study. This includes the four simulated weapons director consoles, the two supporting VAX 780 computers, a 16 channel instrumentation recorder, and two video camera/recorders. Two TV Central camera operators will also be needed.

12. Voluntary Consent Statement: Attached

13. References:

Betts, T., Markman, D., Debenham, S., Mortiboy, D. and McKevitt, T. (1984). Effects of two antihistamine drugs on actual driving performance. British Medical Journal, 288, 281-282.

Boggs, P.B. (1987). Clinical Experience with Terfenadine Worldwide Trials. Proceedings of Current Pharmacotherapy for Allergic Rhinitis and Urticaria. The Journal of Respiratory Disease, 7, 36-46.

Clarke, C.H., and Nicholson, A.N. (1978). Performance studies with antihistamines. British Journal of Clinical Pharmacology, 6, 31-35.

Dyer, J.L. (1986). Annotated Bibliography and state-of-the-art review of the field of team training as it relates to military teams. ARI Research Note 86-18, ARI Field Unit at Fort Benning, Georgia, Fort Benning, Georgia.

Eddy, D.E. (1989). Selected team performance measures in a C^3 Environment--An Annotated Bibliography. Technical Report USAFSAM-TR-87-25, USAF School of Aerospace Medicine, Brooks AFB, Texas.

Englund, E.C., Reeves, R.L., Shingledecker, C.A., Thorne, D.R., Wilson, K.P., and Hegge, F.W. (1985). Unified Tri-Service Cognitive Performance Assessment Battery (UTC-PAB) I: Design and Specification of the Battery, JW6D3 MILPERF Report No. 85-1, U.S. Army Research and Development Command, Fort Detrick, Maryland.

Fink, M., and Irwin, P. (1979). CNS effects of the antihistamines diphenhydramine and terfenadine (RMI 9918). Pharmakopsychiatrist, 12, 35-44.

Kulshrestha, V.K., Gupta, P.P., Turner, P. and Wadsworth, J. (1978). Some clinical pharmacological studies with terfenadine, a new antihistamine drug. British Journal of Clinical Pharmacology, 6, 25- 29.

Nicholson, A.N., Smith, P.A., and Spencer, M.B. (1982). Antihistamines and visual function: Studies on dynamic acuity and the pupillary response to light. British Journal of Clinical Pharmacology, 14, 683-690.

Nicholson, A.N., and Stone, B.M., (1986). Antihistamines: Impaired performance and the tendency to sleep. European Journal of Clinical Pharmacology, 30, 27-32.

Perez, W.A., Masline, P.J., Ramsey, E.G., and Urban, K.E. (1987). Unified tri-services cognitive performance assessment battery: Review and methodology. Technical Report AAMRL-TR-87-007, Armstrong Aerospace Medical Research Laboratory, Wright-Patterson AFB, Ohio.

Samet, M.G., Marshall-Mies, J.C., and Albarian, G. (1987). Expanded complex cognitive assessment battery (CCAB): Final test administrator user guide. AAC-UM-33212, Analytical Assessments Center, CA: Los Angeles. Deliverable under Contract MDA903-84-C-0449 to U.S. Army Research Institute, VA: Alexandria.

Shingledecker, C.A., (1984). A task battery for applied human performance assessment research. Technical Report No. AFAMRL-TR-84-071, Air Force Aerospace Medical Research Laboratories, Wright-Patterson AFB, Ohio.

Sorkin, E.M., and Heel, R.C., (1985). Terfenadine: A review of its pharmacodynamic properties and therapeutic efficacy. International Journal of Current Therapeutics and Applied Pharmacology Reviews, ADIS Press International, Inc. Newtown, PA.

Spector, S.L. (1987). The pharmacology of antihistamines. Proceedings of Current Pharmacotherapy for Allergic Rhinitis and Urticaria. The Journal of Respiratory Diseases, 7, 24-28.

FIGURE 1. Research Design.

Teams of three subjects will receive a different drug treatment, either Benadryl, Seldane, or placebo. Each group will receive its drug treatment on the following schedule.

<u>Benadryl Group</u>

Monday:	Training	- Placebo at 2230
Tuesday:	Testing	- Placebo 3 times, then Benadryl at 2230
Wednesday:	Testing	- Benadryl
Thursday:	Testing	- Benadryl
Friday:	Free, unless needed to make up previous day.	

<u>Seldane Group</u>

Monday:	Training	- Placebo at 2230
Tuesday:	Testing	- Placebo 3 times, then Seldane at 2230
Wednesday:	Testing	- Seldane
Thursday:	Testing	- Seldane
Friday:	Free, unless needed to make up previous day.	

<u>Placebo Group</u>

Monday:	Training	- Placebo at 2230
Tuesday:	Testing	- Placebo
Wednesday:	Testing	- Placebo
Thursday:	Testing	- Placebo
Friday:	Free, unless needed to make up previous day.	

Data collection will normally start no sooner than 1 hour after drug administration.

Subjects will consume no more than six capsules daily. Benadryl subjects will consume no more than 200 mg daily and Seldane subjects will consume no more than 120 mg daily.

FIGURE 2. Complete Design for a Group.

Workload Order 1

	Tues	Wed	Thurs
Drug	P	D	D
Morn	Lo	Lo	Lo
Aft	Hi	Hi	Hi

Workload Order 2

	Tues	Wed	Thurs
Drug	P	D	D
Morn	Hi	Hi	Hi
Aft	Lo	Lo	Lo

NOTES:

1. P - Placebo — Lo - Low-workload session
   D - Drug — Hi - High-workload session

2. In the Placebo Group, the drug would be replaced by placebo.

3. If there should be an equipment/software or subject problem, Friday will be available for testing.

4. Four teams will be divided equally between the two orders.

FIGURE 3. Daily Testing Schedule

Time	Activity
0600	Mission Planning & Breakfast
0630	Drug/Placebo, questionnaires
0700	Pre-brief
0730	Run Scenario
1100	Post-brief, Box lunch
1130	Drug/Placebo, briefing, questionnaires
1230	UTC-PAB
1330	CCAB
1430	Mission Planning & snack
1500	Drug/Placebo, questionnaires
1530	Pre-brief
1600	Run Scenario
1930	Post-brief, questionnaires
2030	Supper, free time (see notes)
2230	Drug/Placebo, but not on Thursday

NOTES: Precautions the night before: Bed early (10:30) light dinner. Keep blood alcohol levels to below legal limits (0.1%).

Breakfast - juice, fruit, toast, water, doughnuts, decaffeinated soft drinks, herbal tea

Lunch - salad, vegetable soup, crackers, cookies, decaffeinated soft drinks, herbal tea, fruit, no chocolate

Voluntary Consent Statement

1.The purpose of this study is to acquire sensitivity and reliability performance data on realistic Air Defense scenario tasks under two antihistamines (drugs for allergies) and a sustained operations schedule. The tasks require differing levels of thought, memory, decision making, and motor response. The drugs are commercially available allergy or antihistamine medications. A secondary purpose of this study is to acquire data relating the actual performance impairment of these antihistamines on these complex, decision-making tasks under normal, extended-day duty conditions.

2.I will complete one full training day and three 15 1/2 hr testing sessions (see Figure 1). At the end of training on Monday, I will ingest the first dose of either Seldane (60mg), Benadryl (50mg), or a placebo. The placebo is a non-active lactose sugar pill. I understand that if I have lactose sensitivity or am allergic to milk, I should not take part in this study. I understand that my knowledge or the investigators' knowledge of my medication treatment may affect my performance. Therefore, neither I nor the investigators will know which medication (antihistamine or placebo) I am taking. Only the medical monitor will know that information. I understand that I will be taking no more than the recommended therapeutic regimen (four 50mg Benadryl pills, or two 60mg Seldane pills) or the same number of placebo pills during each 15 1/2 hr session. As an additional protection against identifying which medication I will be given, I will receive the same number of capsules per day (6) as the total of both drugs, although the doses may be either all placebo or placebos plus daily doses of one medication.

3.The attached schedule shows (Figure 1) the times that I will take the medication and when I will fill out questionnaires and perform the tasks. Breakfast and lunch will be provided at the times shown on the schedule. I will drink only caffeine-free soft drinks and will smoke only in designated outside areas during off-task time.

4.By taking a recommended therapeutic dose of either antihistamine during a 24-hr period, I may experience some discomfort related to some common side effects. The side effects may include drowsiness, headache, nausea, or dryness of the mouth/nose/throat. I understand that the potential for experiencing these side effects is low and by completing the symptoms questionnaire in the morning, afternoon, and evening, I will give the investigators information concerning the frequency of such side effects to keep the medical monitor informed.

____

Volunteer Signature Date

5.If female, I have had a negative pregnancy test within 30 days prior to ingesting the initial dose. I understand that I will not be allowed to participate in this study if I am pregnant. I have read and signed the attached pregnancy disclaimer. I understand that experience with these drugs in pregnant women is inadequate to determine whether there exists a potential for harm to the developing fetus. I do not believe I am pregnant and have signed the attached pregnancy disclaimer.

6.I will benefit from participation in this study by knowing that I have contributed data to an important crew performance data base.

7.There is no other alternative procedure that will more effectively acquire this important information.

8.Records of my participation in this study may only be disclosed according to federal law, including the Federal Privacy Act, 5 U.S.C. 552a, and its implementing regulations.

9.I understand that my entitlement to medical care or compensation in the event of injury are governed by federal laws and regulations, and if I desire further information, I may contact the Chief of Flight Medicine (536-2859), the Base Legal Office (536-3301), or the principal/associate investigators (532-5723; 536-3464).

10.The decision to participate in this research is completely voluntary on my part. No one has coerced or intimidated me into participating in this program. I am participating because I want to. Dr. Eddy has adequately answered any and all questions I have about this study, my participation, and the procedures involved. I understand that Dr. Eddy, Dr. Schiflett, or Dr. French will be available to answer any questions concerning procedures throughout this study. I understand that if significant new findings develop during the course of this research that may relate to my decision to continue participation, I will be informed. I further understand that I may withdraw this consent at any time and discontinue further participation in this study without prejudice to my entitlements. I also understand that the medical monitor of this study may terminate my participation in this study if he or she feels this to be in my best interest.

Volunteer Signature and Social Security Number Date

Witness (not directly involved) Date

Witness (not directly involved) Date

Pregnancy Disclaimer for Female Subjects

I understand that experience with terfenadine (Seldane) and diphenhydramine (Benadryl) in pregnant women is inadequate to determine whether there exists a potential for harm to the developing fetus. I have submitted to a pregnancy test within 30 days of starting this study. My test was negative.

However, in the event that I am or become pregnant, I relieve the Air Force of any and all responsibility for medical complications, associated with such a pregnancy, that may result from the ingestion of either antihistamine during this study.

____

Volunteer Signature   Date

The pregnancy test for___________________________ is negative.

Name

____

Physician   Date

APPENDIX B

SUBJECT DATA

PARTICIPANT INFORMATION

Subject Number	Male/ Female	Age (Yrs)	Rank	Total E-3 Hours	Total Sim Hours
01	M	27	1Lt	680	1000
02	M	24	1Lt	500	750
07	M	24	1Lt	280	200
08	M	24	1Lt	285	200
11	F	25	1Lt	175	350
12	M	28	1Lt	1300	700
13	M	27	1Lt	90	175
14	M	24	1Lt	145	166
19	M	27	Cpt	510	1000
20	M	25	1Lt	1500	500
23	M	30	1Lt	950	250
24	M	30	1Lt	540	170
29	F	27	1Lt	1000	300
30	M	29	1Lt	215	320
35	F	23	2Lt	100	170
36	M	25	1Lt	315	230

APPENDIX C

SPEECH ANALYSIS SUMMARY DATA

2OCT/WD1

DURATION	FUNDAMENTAL FREQUENCY (Hz)	JITTER (%)	SHIMMER (%)
55.43	179.9	0.47	4.29
41.39	187.36	0.8	10.01
78.47	200.55	0.59	12.53
97.79	219.75	0.55	3.47
119.93	239.86	0.54	8.12
96.1	198.26	0.47	5.42
79.07	189.52	0.54	8.0
WORKLOAD			
10	179.9	0.47	4.29
60	187.36	0.8	10.01
130	200.55	0.59	12.53
145	210.82	0.3	7.05
135	217.43	0.47	7.47

2OCT/WD2

DURATION	FUNDAMENTAL FREQUENCY (Hz)	JITTER (%)	SHIMMER (%)
50.48	125.11	0.63	8.76
98.3	134.19	0.57	16.09
91.36	130.51	0.62	12.89
98.39	163.88	0.72	11.93
107.74	170.23	0.61	4.61
115.62	166.64	0.98	7.32
46.79	138.54	0.38	4.25
WORKLOAD			
10	125.11	0.63	8.76
60	142.67	1.5	4.56
130	141.32	0.67	7.13
145	163.4	0.8	11.94
135	130.51	0.84	8.61

7AUG/WD1

DURATION	FUNDAMENTAL FREQUENCY (Hz)	JITTER (%)	SHIMMER (%)
6.1	104.24	1.11	10.56
105.3	121.12	0.53	9.12
95.95	152.66	0.47	4.55
94.58	169.69	0.91	5.22
97.48	127.99	0.75	5.93
108.95	142.33	0.99	5.36
30.94	126.87	0.8	24.68
WORKLOAD			
10	104.24	1.11	10.56
60	149.9	0.58	3.08
130	152.66	0.47	4.55
145	169.69	0.91	5.22
135	119.11	0.56	6.26

7AUG/WD2

DURATION	FUNDAMENTAL FREQUENCY (Hz)	JITTER (%)	SHIMMER (%)
8.25	136.99	1.22	12.18
62.6	170.26	0.61	14.85
89.71	164.14	0.58	11.3
61.29	178.05	0.72	15.02
76.57	148.38	0.34	9.09
59.64	168.35	1.83	15.59
11.5	148.53	1.17	17.18
WORKLOAD			
10	136.99	1.22	12.18
60	155.31	1.17	13.44
130	157.84	1.14	13.0
145	168.69	1.69	13.24
135	148.38	0.34	9.09

10JUL/WD1

DURATION	FUNDAMENTAL FREQUENCY (Hz)	JITTER (%)	SHIMMER (%)
58.63	129.55	0.95	7.46
54.65	177.06	0.53	5.75
103.28	164.27	0.81	10.32
93.64	164.30	0.81	15.34
86.28	167.20	0.80	6.74
75.0	147.62	0.85	5.61
61.9	169.53	1.71	8.65
WORKLOAD			
10	129.83	0.96	6.90
60	135.15	2.33	15.1
130	164.27	0.81	10.32
145	153.73	0.94	11.49
135	167.54	0.44	6.05

10JUL/WD2

DURATION	FUNDAMENTAL FREQUENCY (Hz)	JITTER (%)	SHIMMER (%)
33.16	68.44	5.28	12.52
48.03	87.53	1.10	11.45
64.39	147.62	1.25	5.02
65.41	133.35	0.33	6.40
47.81	149.86	1.25	8.34
31.02	115.08	0.81	6.81
21.92	140.89	1.06	7.84
WORKLOAD			
10	99.2	0.70	15.53
60	95.19	2.63	9.73
130	147.59	1.27	4.71
145	143.12	0.58	12.53
135	142.94	1.44	11.51

11SEP/WD1

DURATION	FUNDAMENTAL FREQUENCY (Hz)	JITTER (%)	SHIMMER (%)
3.97	101.74	0.51	7.30
74.43	107.57	1.01	14.75
80.0	113.88	0.67	16.29
110.32	104.51	1.16	16.33
103.17	123.50	0.68	11.58
75.97	120.83	0.49	7.53
26.25	106.63	0.56	10.75
WORKLOAD			
10	101.74	0.51	7.30
60	97.19	0.67	14.22
130	113.88	0.67	16.29
145	104.51	1.16	16.33
135	115.38	0.69	17.25

11SEP/WD2

DURATION	FUNDAMENTAL FREQUENCY (Hz)	JITTER (%)	SHIMMER (%)
58.27	111.74	3.12	8.88
89.51	126.05	1.13	9.93
122.14	138.17	0.62	9.58
92.88	126.93	0.63	10.94
128.25	166.93	1.31	8.56
98.04	130.17	0.98	12.74
75.65	145.75	0.54	5.50
WORKLOAD			
10	111.74	3.12	8.88
60	124.26	0.48	11.34
130	125.33	0.85	10.8
145	126.93	0.63	10.94
135	166.93	1.31	8.56

14AUG/WD1

DURATION	FUNDAMENTAL FREQUENCY (Hz)	JITTER (%)	SHIMMER (%)
4.71	186.80	0.60	11.67
41.89	238.67	0.58	5.97
61.48	199.96	1.65	5.29
43.11	227.41	0.89	8.70
111.12	214.44	0.92	6.06
98.61	242.87	0.68	6.82
51.18	222.01	0.76	5.21
WORKLOAD			
10	186.80	0.60	11.67
60	238.67	0.58	5.97
130	235.94	0.46	6.44
145	227.41	0.89	8.70
135	210.93	1.14	8.79

14AUG/WD2

DURATION	FUNDAMENTAL FREQUENCY (Hz)	JITTER (%)	SHIMMER (%)
13.25	106.26	0.61	8.15
111.88	132.44	0.78	15.39
147.03	145.71	0.69	14.54
111.75	141.85	0.71	14.75
117.46	129.77	0.54	13.1
77.3	129.82	0.41	7.4
47.25	131.04	0.56	13.5
WORKLOAD			
10	106.26	0.61	8.15
60	132.44	0.78	15.39
130	121.4	0.53	9.32
145	141.85	0.71	14.75
135	129.77	0.54	13.1

16OCT/WD1

DURATION	FUNDAMENTAL FREQUENCY (Hz)	JITTER (%)	SHIMMER (%)
9.87	189.31	0.82	5.69
50.61	182.67	1.28	9.46
83.72	173.36	2.67	10.35
99.74	196.4	0.74	9.03
97.65	228.44	1.28	10.05
99.71	252.82	0.69	10.61
71.34	216.32	1.58	11.94
WORKLOAD			
10	189.31	0.82	5.69
60	193.27	1.83	9.47
130	173.36	2.67	10.35
145	178.68	0.72	7.74
135	182.35	0.45	6.47

16OCT/WD2

DURATION	FUNDAMENTAL FREQUENCY (Hz)	JITTER (%)	SHIMMER (%)
53.22	120.17	1.04	14.37
49.65	106.77	0.41	11.18
48.42	109.30	1.23	11.70
72.40	144.14	0.90	6.12
68.19	132.32	0.63	6.37
74.37	153.06	1.53	7.04
40.4	115.77	1.45	7.69
WORKLOAD			
10	120.17	1.04	14.37
60	111.6	0.87	8.75
130	109.3	1.23	11.7
145	119.25	1.95	12.4
135	132.32	0.63	6.37

18SEP/WD1

DURATION	FUNDAMENTAL FREQUENCY (Hz)	JITTER (%)	SHIMMER (%)
7.78	74.42	5.48	15.87
73.44	97.69	2.11	15.48
76.43	104.15	0.88	12.63
113.21	109.69	0.92	8.37
80.23	140.82	0.45	8.27
94.63	143.81	0.83	8.58
52.05	109.14	1.80	13.83
WORKLOAD			
10	74.41	5.52	16.02
60	105.82	1.64	10.43
130	106.89	1.38	13.63
145	98.79	0.67	8.20
135	136.25	1.27	10.0

18SEP/WD2

DURATION	FUNDAMENTAL FREQUENCY (Hz)	JITTER (%)	SHIMMER (%)
16.25	93.31	8.97	7.26
44.59	117.46	1.64	16.61
93.66	130.62	1.58	15.47
63.29	113.28	0.53	5.85
85.68	137.87	2.37	18.6
67.20	147.85	0.99	9.24
47.77	108.26	1.60	17.31
WORKLOAD			
10	93.64	7.72	10.19
60	116.94	1.69	17.95
130	113.03	1.33	16.40
145	110.73	1.30	16.26
135	143.29	1.03	14.01

21AUG/WD1

DURATION	FUNDAMENTAL FREQUENCY (Hz)	JITTER (%)	SHIMMER (%)
40.49	169.11	1.09	11.42
53.61	152.37	1.21	11.03
74.43	187.84	1.09	7.42
74.77	169.55	0.87	7.09
76.11	180.77	0.72	7.97
89.11	197.59	1.10	5.99
57.68	203.97	0.78	4.65
WORKLOAD			
10	169.11	1.09	11.42
60	152.37	1.21	11.03
130	187.84	1.09	7.42
145	181.51	1.07	7.38
135	155.45	1.25	17.13

21AUG/WD2

DURATION	FUNDAMENTAL FREQUENCY (Hz)	JITTER (%)	SHIMMER (%)
16.94	144.41	0.64	7.63
116.26	145.89	0.61	6.61
125.84	181.61	0.71	6.54
89.36	177.79	0.47	6.29
111.79	195.02	0.99	5.43
84.08	183.18	0.28	5.05
25.84	148.93	0.57	5.29
WORKLOAD			
10	144.41	0.64	7.63
60	172.04	0.35	11.13
130	164.61	0.35	3.95
145	177.79	0.47	6.29
135	158.46	0.34	10.89

APPENDIX D

STATISTICAL RESULTS FROM ANOVA, REGRESSION, AND CORRELATION DATA

ANOVA TABLES

MODEL.FREQUENCY

Source	df	SS	MS	F	P
Subject	15	81594.009	5439.6006	26.96	<.0001
Levels	4	6157.7187	1539.4297	7.63	<.0001 *
Error	60	12105.933	201.76555		

MODEL.JITTER

Source	df	SS	MS	F	P
Subject	15	10.694755	.71298367	1.96	.0338
Levels	4	2.2259550	.55648875	1.53	.2040
Error	60	21.774445	.36290742		

MODEL.SHIMMER

Source	df	SS	MS	F	P
Subject	15	447.23746	.29815831	3.24	.0006
Levels	4	15.036670	3.7591675	.41	.8021
Error	60	552.71289	9.2118815		

DUNCAN'S MULTIPLE RANGE TEST FOR VARIABLE: FREQUENCY

DUNCAN GROUPING @

.05	.01	.001	Mean	N	Workload
A	A	A	154.806	16	H2
A	A	A	153.565	16	H1
A	A	A	150.988	16	Med
A	A	A B	144.386	16	Low
B	B	B	129.554	16	Baseline

RELATIONSHIP OF MODEL WITH FREQUENCY

SUBJECTS	N	SLOPE	P-VALUE	R-VALUE
2OCT/WD1	5	.247806	.0229	.928
2OCT/WD2	5	.145664	.3046	.581
7AUG/WD1	5	.275200	.2799	.605
7AUG/WD2	5	.148649	.1505	.743
10JUL/WD1	5	.262980	.0325	.909
10JUL/WD2	5	.410135	.0249	.924
11SEP/WD1	5	.090359	.2110	.675
11SEP/WD2	5	.211188	.2936	.591
14AUG/WD1	5	.198577	.3444	.543
14AUG/WD2	5	.162170	.1785	.711
16OCT/WD1	5	-.107275	.1150	-.786
16OCT/WD2	5	.026703	.7803	.173
18SEP/WD1	5	.104906	.5068	.398
18SEP/WD2	5	.205509	.2420	.643
21AUG/WD1	5	.095448	.5525	.359
21AUG/WD2	5	.136216	.2662	.618
POOLED DATA	80	.163390	.0001	.616

RELATIONSHIP OF MODEL WITH JITTER

SUBJECTS	N	SLOPE	P-VALUE	R-VALUE
2OCT/WD1	5	-.00117139	.5375	-.372
2OCT/WD2	5	-.00097604	.7944	-.162
7AUG/WD1	5	-.00246405	.3564	-.532
7AUG/WD2	5	-.00088671	.8640	-.107
10JUL/WD1	5	-.00549601	.4499	-.447
10JUL/WD2	5	-.00196950	.8203	-.142
11SEP/WD1	5	.00279956	.2175	.668
11SEP/WD2	5	-.01268990	.1957	-.692
14AUG/WD1	5	.00226797	.4119	.481
14AUG/WD2	5	-.00048439	.6711	-.261
16OCT/WD1	5	.00052360	.9579	.033
16OCT/WD2	5	.00299056	.5640	.350
18SEP/WD1	5	-.02006830	.0333	-.908
18SEP/WD2	5	-.00758097	.0123	-.953
21AUG/WD1	5	.00006826	.9378	[illegible]49
21SEP/WD2	5	-.00141249	.2426	-.642
POOLED DATA	80	-.00290936	.0245	-.323

RELATIONSHIP OF MODEL WITH SHIMMER

SUBJECTS	N	SLOPE	P-VALUE	R-VALUE
2OCT/WD1	5	.02421930	.4422	.454
2OCT/WD2	5	.01784680	.5166	.390
7AUG/WD1	5	-.02646480	.3386	-.548
7AUG/WD2	5	-.00624909	.7422	-.204
10JUL/WD1	5	.00093246	.9810	.015
10JUL/WD2	5	-.03361370	.3991	-.493
11SEP/WD1	5	.06441980	.0211	.932
11SEP/WD2	5	.00473348	.7267	.216
14AUG/WD1	5	-.01673710	.4673	-.432
14AUG/WD2	5	.02146620	.5188	.388
16OCT/WD1	5	.01112780	.5844	.333
16OCT/WD2	5	-.02156790	.5039	-.401
18SEP/WD1	5	.00680465	.7546	.194
18SEP/WD2	5	.02184390	.4123	.481
21AUG/WD1	5	-.00706100	.8680	-.104
21AUG/WD2	5	-.01377200	.6684	-.264
POOLED DATA	80	.00299555	.6406	.103

RELATIONSHIP OF DURATION WITH FREQUENCY

SUBJECTS	N	SLOPE	P-VALUE	R-VALUE
2OCT/WD1	9	.637035	.0122	.785
2OCT/WD2	10	.411485	.0582	.615
7AUG/WD1	9	.334316	.0850	.604
7AUG/WD2	10	.224678	.1611	.479
10JUL/WD1	10	.347294	.2550	.398
10JUL/WD2	11	.713576	.2381	.388
11SEP/WD1	9	.139764	.1258	.549
11SEP/WD2	9	.442458	.0697	.629
14AUG/WD1	9	.239264	.2853	.401
14AUG/WD2	8	.218852	.0229	.778
16OCT/WD1	10	.325710	.2876	.374
16OCT/WD2	9	1.217140	.0006	.912
18SEP/WD1	11	.150450	.4596	.249
18SEP/WD2	10	.623172	.0093	.769
21AUG/WD1	9	.618565	.1125	.566
21AUG/WD2	10	.286219	.0631	.606
POOLED DATA	153	.271899	.0001	.591

RELATIONSHIP OF DURATION WITH JITTER

SUBJECTS	N	SLOPE	P-VALUE	R-VALUE
2OCT/WD1	9	-.00216285	.3123	-.381
2OCT/WD2	10	.00352376	.4285	.283
7AUG/WD1	9	-.00314425	.1594	-.511
7AUG/WD2	10	-.00944008	.1223	-.521
10JUL/WD1	10	-.01607730	.1374	-.504
10JUL/WD2	11	-.02431560	.4208	-.271
11SEP/WD1	9	.00393017	.0852	.604
11SEP/WD2	9	-.01622520	.2262	-.448
14AUG/WD1	9	.00228537	.6111	-.582
14AUG/WD2	8	.00092670	.4147	.337
16OCT/WD1	10	-.00232595	.7831	-.100
16OCT/WD2	9	.00000335	.9998	0.000
18SEP/WD1	11	-.02977910	.0005	-.870
18SEP/WD2	10	-.00529203	.5609	-.210
21AUG/WD1	9	-.00405063	.3550	-.351
21AUG/WD2	10	.00156169	.4596	.265
POOLED DATA	153	-.00613988	.0171	-.214

RELATIONSHIP OF DURATION WITH SHIMMER

SUBJECTS	N	SLOPE	P-VALUE	R-VALUE
2OCT/WD1	9	-.0246728	.5970	-.205
2OCT/WD2	10	.0473866	.4065	.296
7AUG/WD1	9	-.1217910	.0431	-.682
7AUG/WD2	10	-.0370742	.2177	-.428
10JUL/WD1	10	.0108384	.8853	.053
10JUL/WD2	11	-.1034700	.1170	-.500
11SEP/WD1	9	.0554770	.1610	.510
11SEP/WD2	9	.0186814	.5987	.204
14AUG/WD1	9	-.0393599	.1000	-.582
14AUG/WD2	8	.0428311	.1351	.576
16OCT/WD1	10	.0290656	.2041	.439
16OCT/WD2	9	-.1202780	.1811	-.490
18SEP/WD1	11	-.0276229	.3642	-.304
18SEP/WD2	10	.0331399	.6182	.180
21AUG/WD1	9	-.1607710	.0407	-.687
21AUG/WD2	10	.0048519	.8367	.075
POOLED DATA	153	-.0164525	.3424	-.124

VITA

Name:	William Archer Jones, Jr.
Date of Birth:	April 26, 1963
Place of Birth:	Bloomfield, Connecticut
Permanent Address:	14 Candlewood Drive Bloomfield, Connecticut 06002
Education:	Bachelor of Science, 1985 United States Air Force Academy Colorado Springs, Colorado Major: Human Factors Engineering
	Master of Science, 1990 Texas A&M University College Station, Texas Major: Industrial Engineering
Organizations/Societies:	Human Factors Society Tau Beta Pi Honor Society Alpha Pi Mu Honor Society
Profession:	Captain, United States Air Force
Experience:	Foreign Weapons/Human Factors Analyst Foreign Technology Division Wright-Patterson AFB, Ohio June, 1985 to August, 1988